

An Electronic PDR™ Using Fielded Searches in HTML

Jeff Zucker, M.A., Robert M. Kahn, Ph.D., Pat Molholt, M.L.S.

Office of Scholarly Resources, Health Sciences Division

Columbia University, New York, NY

At Columbia-Presbyterian Medical Center (CPMC), we have developed a version of Physician's Desk Reference™ (PDR) using HyperText Markup Language, or HTML, the standard markup language of the World-Wide Web. It is accessible on the campus network and from anywhere on the Internet for staff and students with appropriate access codes. We have previously reported^{1,2} earlier electronic versions of the PDR using Folio Views™ software, and our efforts toward creating an integrated information environment at CPMC. The development of an HTML version of the PDR is a logical extension of these previous projects, and takes advantage of ubiquitous HTML browsers such as Netscape, Mosaic, and Lynx to present drug information in a way that is familiar to an ever increasing number of users.

Three times a year we receive tapes of the full text of the PDR for Prescription Drugs™, the PDR for Non-Prescription Drugs™, and the PDR for Ophthalmologic Drugs™ from the publisher, Medical Economics Data. Using the Perl language, we parse the text into two formats: i) an array of HTML files and ii) a database that contains fields for trade name, generic class, therapeutic class, and pointers to the HTML files for each product. These files, along with additional Perl scripts based on the Common Gateway Interface (CGI) to HTML, form the basis of our online version.

The user's first entry to the online PDR is a single line form into which can be typed a trade name (e.g., Biaxin), a generic class (e.g., clarithromycin), or a therapeutic class (e.g., antibiotics). The search is based on substrings so 'biax', 'clarithro', and 'antib' will work just as well. The search also allows two word combinations in any order. Thus 'potassium penicillin' and 'penicillin potassium' are treated the same. This approach was based in part on feedback which showed that many users knew only part of the name of a drug, and also did not know if the name they had in mind was a generic or trade name.

Once the search phrase is obtained, the script matches it against the pre-built fielded database instead of against the array of HTML files, which has many advantages. For example, it allows a much more rapid search, since the database is approximately 150 kilobytes, compared to the full text, which is over 50 megabytes. Retrieval of information using the fielded database also yields greater

precision than a full text search such as the Wide Area Information Server (WAIS) search engine. Finally, a fielded search allows a more informative display of the matches; users are presented first with the most specific responses (i.e., matching trade names) that then widen out to more general (e.g., generic or therapeutic class) matches.

When the user selects a trade name from the search results display, they are presented with the full text of the PDR reference to that product. The user may then scroll through or use a set of internal anchors. These anchors, built by the original Perl script, allow jumps to specific areas within the text, such as to dosage information. Anchors are specific to each drug; thus, only those drugs that contain sections on animal pharmacology show 'active' anchors to an animal pharmacology section. Several other options are also presented; hypertext buttons at the top of the file allow the user to i) enter another search, ii) move to an alphabetic index of all drug names, iii) get 'help', iv) send 'feedback', or v) view 'related drugs', which pulls up a list of all the generic and therapeutic classes that contain the current drug, allowing the user to jump to other drugs related along several axes of similarity.

References

1. Zucker J, Kahn RM, Narayanan N. Clinical, scholarly, and campus information hypertext tools at CPMC. Proc 17th Annu Symp Comput Appl Med Care. 1993; 539-543.
2. Kahn RM, Molholt P, Zucker J. CPMCnet: An Integrated Information and Internet Resource. Proc 18th Annu Symp Comput Appl Med Care. 1994; 98-102.

Acknowledgement: This publication was developed under the auspices of the Columbia University CAT in High Performance Computing and Communications in Healthcare, a New York State Center for Advanced Technology supported by the New York State Science and Technology Foundation.